Harvard University PhD Program in Health Policy
Evaluative Science and Statistics Concentration
2017-2018

TABLE OF CONTENTS
• INTRODUCTION
• REQUIRED COURSES
• STATISTICS COURSE REQUIREMENTS
• EVALUATIVE SCIENCES COURSE REQUIREMENTS
• SUGGESTED SEQUENCE OF STATISTICS COURSES
• LIST OF CLASSES TAKEN BY CURRENT STUDENTS
• FACULTY MEMBERS
• COURSE DESCRIPTIONS

INTRODUCTION
General description of the evaluative science concentration:

Training in this concentration will enable students to study the effects of a wide range of policies and health services (e.g., health insurance, health-care quality improvement, clinical decision-making, drug policy, cost-containment, and socioeconomic factors) on behaviors, access, quality of care, health outcomes, and costs. Students in this concentration will develop proficiency in experimental and quasi-experimental research design, statistics, relevant social sciences, and other methodological approaches (e.g., epidemiology, program evaluation, qualitative methods, and survey design). Previous students in this track have used innovative methodological and statistical approaches to study, for example:

POLICY EVALUATION
• Impact of acquiring Medicare coverage on the health of previously uninsured adults
• Effects on health behavior of insurance restrictions on maternity lengths of stay
• Effects of drug coverage on access to essential medications in Medicare
• Effects of regulatory changes in legal drinking ages on health and mortality

QUALITY OF CARE AND CLINICAL DECISION-MAKING
• Effects of physician experiences with adverse medical events on under-prescribing of essential medicines
• A controlled natural experiment on the effectiveness of direct to consumer drug advertising

DISPARITIES
• Effects of near-universal Medicare coverage on disparities in cardiovascular disease and diabetes control
• Methods to estimate racial/ethnic health care disparities and their effects on health

COMPARATIVE HEALTH POLICY
• International differences in health outcomes following medical care for acute myocardial infarction

POPULATION HEALTH
• Effects of unemployment on mortality

METHODS
• Potential bias of instrumental variable analyses for observational comparative effectiveness research

Click to view descriptions of dissertation research for all graduates of the Evaluative Science and Statistics track.
Courses for students in the evaluative science concentration:
The course requirements for the evaluative science and statistics concentration were selected to provide students with important skills needed for conducting original health policy research; at the completion of coursework, students should be able to propose feasible study designs to answer health policy questions, using both experimental and quasi-experimental designs, as well as identify the strengths and limitations of the various designs in their proposed work and in other published studies. They also will develop strong analytical skills, including the technical expertise required to analyze data as well as interpret results, identify the strengths and limitations of analyses, and the broader implications of results for future health policy.

The following course requirements encompass both statistics and research design. Students are required to take at least 5.0 credits of statistics; students should be proficient in linear, logistic and survival regression analyses upon completion of statistics-related coursework. Students are also strongly recommended to develop skills in hierarchical modeling and analysis of survey data, either through coursework or outside study. Incoming students should also be aware that many of the intermediate and advanced courses have prerequisites. In addition, some courses are offered only in alternative academic years, and scheduling conflicts do occur. Due to these issues, incoming students should plan a tentative two-year course program early in the first year. They should also consult with more senior students about their tentative program for additional feedback and suggestions. In planning such a program, students are encouraged to follow a basic sequence of courses within a particular school or department (i.e., Biostatistics, Statistics, Economics, or HKS) as opposed to selecting from multiple schools/departments. This is recommended in order to facilitate a more coherent presentation of the fundamentals. Subsequently, students are encouraged to choose courses based on interests and career objectives from among all schools/departments.

REQUIRED COURSES
(Note: 1 credit = 1 semester):

- **Statistics (5.0 credits)**
  - Probability Theory (1.0 credit)
  - Statistical Inference (1.0 credit)
  - Regression (1.0 credit)
  - Other Data Analyses (2.0 credits)

- **Evaluative Sciences (4.0 credits)**
  - Research Design and Methods (2.0 credits)
    - GHP 228 (1.0 credit)
    - Health Policy 3080A&B (1.0 credits)
  - Specific Methodological Approaches (2.0 credits)
    - Survey Research Methods (recommended)
    - Causal Inference (recommended)
    - Decision Sciences
    - Epidemiology
    - Program Evaluation
    - Qualitative Research
    - Network Analysis
    - Other

- **Program and Distribution Requirements (6.5 credits)**
  - Health Policy Core Course (2.0 credits)
  - 3rd/4th year health policy seminar (1.0 credit)
  - Distribution requirements (3.0 credits; 1.0 credit in three of the following areas)
    - Decision Sciences
    - Economics
    - Ethics
    - Management
    - Political Analysis

Notes:
1) Students in the ESS track are strongly encouraged to take courses in Survey Research Methods and Causal Inference. Students are also encouraged to supplement their curriculum with social sciences classes to gain a well-rounded understanding of the health care system. Finally, students should make an effort to regularly attend research design seminars that are held around the university.
2) ESS students may count decision sciences courses towards the PhD distribution requirements as well as towards the Specific Methodological Approaches requirement.
3) Certain courses are listed twice under different requirements, but they CANNOT count toward more than one requirement (e.g., BST 222 can count towards either Probability Theory or Statistical Inference, but not both). The classes that are listed under two requirements but that only count toward one requirement are:

- BST 222 and Econ 2110 count as a Probability Theory or Statistical Inference course
- Gov 2001 counts as a Statistical Inference or Other Data Analysis course
- BST 232 counts as a Regression or Other Data Analysis course
- Statistics 160 counts as an Other Data Analysis or Survey Research Methods course
- Statistics 186 counts as a Statistical Inference or Causal Inference course

4) To satisfy all ESS, program, and distribution requirements, ESS students should plan on taking 3-4 credits each semester in their first and second years. Many students TF in their second/third years, and some students take courses in their third year.

5) Students may petition the ESS concentration chairs for credit if they have previously taken one of the courses listed in this packet.

6) Students who have taken a class similar to Statistics 110 should contact Alan Zaslavsky. These students may take a more advanced probability course to satisfy the Probability Theory requirement. Alternatively, they may take an additional statistics course instead of another probability course to satisfy the Probability Theory requirement.

7) Students may petition the ESS concentration chairs if they would like a non-listed course to count towards a requirement. They may petition the DGS if they would like a non-listed course to count towards a distribution requirement.
## STATISTICS COURSE REQUIREMENTS

<table>
<thead>
<tr>
<th>Class</th>
<th>Semester</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Probability Theory (1.0 credit)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistics 110: Introduction to Probability</td>
<td>Fall</td>
<td>TuTh 2:30–4</td>
</tr>
<tr>
<td>Statistics 171(^1): Introduction to Stochastic Processes</td>
<td>Spring</td>
<td>MW 2:30–4</td>
</tr>
<tr>
<td>Statistics 210(^2): Probability I</td>
<td>Fall</td>
<td>MW 11:30–1</td>
</tr>
<tr>
<td>API 201 (A,B): Quantitative Analysis and Empirical Methods</td>
<td>Fall</td>
<td>TTh 8:45–10</td>
</tr>
<tr>
<td>BST 222: Basics of Statistical Inference</td>
<td>Fall</td>
<td>TBA</td>
</tr>
<tr>
<td>Biostatistics 230 / BST 230: Probability Theory and Applications I</td>
<td>Fall</td>
<td>MW 2–3:30</td>
</tr>
<tr>
<td><strong>Statistical Inference (1.0 credit)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistics 111: Introduction to Theoretical Statistics</td>
<td>Spring</td>
<td></td>
</tr>
<tr>
<td>Statistics 211: Statistical Inference I</td>
<td>Spring</td>
<td>WF 2:30–4</td>
</tr>
<tr>
<td>API 201 (A,B): Quantitative Analysis and Empirical Methods</td>
<td>Fall</td>
<td>TTh 8:45–10</td>
</tr>
<tr>
<td>Government 2001: Advanced Quantitative Research Methodology</td>
<td>Spring</td>
<td>M 2–4</td>
</tr>
<tr>
<td>BST 222: Basics of Statistical Inference</td>
<td>Fall</td>
<td>TBA</td>
</tr>
<tr>
<td>Biostatistics 231 / BST 231(^3): Statistical Inference I</td>
<td>Spring</td>
<td>MW 9:45–11:15</td>
</tr>
<tr>
<td>API 209: Advanced Quantitative Methods I: Statistics</td>
<td>Fall</td>
<td>TuTh 11:45–1</td>
</tr>
<tr>
<td><strong>Regression (1.0 credit)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistics 139: Statistical Sleuthing Through Linear Models</td>
<td>Fall</td>
<td>MW 9:30–11</td>
</tr>
<tr>
<td>Statistics 244(^4): Linear and Generalized Linear Models</td>
<td>Fall</td>
<td>TuTh 2:30–4</td>
</tr>
<tr>
<td>Economics 1123: Introduction to Econometrics</td>
<td>Fall/Spring</td>
<td>(F) TuTh 11:30–1; (S) MW 2:30–4</td>
</tr>
<tr>
<td>API 201 (C,D): Quantitative Analysis and Empirical Methods</td>
<td>Spring</td>
<td>TTh 8:45–10</td>
</tr>
<tr>
<td>Government 2000: Introduction to Quantitative Methods I</td>
<td>Fall</td>
<td>M 10–12</td>
</tr>
<tr>
<td>BST 211: Regression &amp; Analysis of Variance in Experimental Research</td>
<td>NA</td>
<td>Not offered 2017–18</td>
</tr>
<tr>
<td>BST 213: Applied Regression for Clinical Research</td>
<td>Fall</td>
<td>MW 8–9:30</td>
</tr>
<tr>
<td>Biostatistics 232/ BST 232: Methods I</td>
<td>Fall</td>
<td>MW 9:45–11:15</td>
</tr>
<tr>
<td>Biostatistics 235 / BST 235(^3): Advanced Regression and Statistical Learning</td>
<td>Fall</td>
<td>MW 9:45–11:15</td>
</tr>
<tr>
<td>GHP 325: Econometrics for Health Policy</td>
<td>Fall</td>
<td>TuTh 8–9:30</td>
</tr>
<tr>
<td>API- 10: Advanced Quantitative Methods II: Econometric Methods</td>
<td>Spring</td>
<td>TuTh 10:15–11:30; review F</td>
</tr>
<tr>
<td><strong>Other Data Analyses (2.0 credits)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistics 120: Introduction to Bayesian Inference and Applications</td>
<td>Fall</td>
<td>MW 9:30–11</td>
</tr>
<tr>
<td>Statistics 131: Time Series and Prediction</td>
<td>Fall</td>
<td>TuTh 1–2:30</td>
</tr>
<tr>
<td>Statistics 149: Statistical Sleuthing Through Generalized Linear Models</td>
<td>Spring</td>
<td>MW 8:30–10</td>
</tr>
<tr>
<td>Statistics 151: Multilevel and Longitudinal Models</td>
<td>NA</td>
<td>Not offered 2017–18</td>
</tr>
<tr>
<td>Statistics 160/260: Design and Analysis of Sample Surveys</td>
<td>Fall</td>
<td>MW 2:30–4</td>
</tr>
<tr>
<td>Statistics 220(^5): Bayesian Data Analysis</td>
<td>Spring</td>
<td>TTh 1–2:30</td>
</tr>
<tr>
<td>Statistics 232R: Topics in Missing Data</td>
<td>NA</td>
<td>Not offered 2017–18</td>
</tr>
<tr>
<td>Economics 2140(^6): Econometric Methods</td>
<td>Spring</td>
<td>TuTh 1:30–3</td>
</tr>
<tr>
<td>Government 2001: Advanced Quantitative Research Methodology</td>
<td>Spring</td>
<td>M 2–4</td>
</tr>
<tr>
<td>Government 2735: Empirical Models in International Relations</td>
<td>NA</td>
<td>Not offered 2017–18</td>
</tr>
<tr>
<td>Sociology 211 (formerly 203b): Analysis of Longitudinal Data</td>
<td>NA</td>
<td>Not offered 2017–18</td>
</tr>
<tr>
<td>BST 210: Applied Regression Analysis</td>
<td>Fall/Spring</td>
<td>(F) MW 8–9:30; (S) TuTh 8–9:30;</td>
</tr>
<tr>
<td>BST 223: Applied Survival Analysis</td>
<td>Spring</td>
<td>TuTh 9:45–11:15</td>
</tr>
<tr>
<td>BST 226: Applied Longitudinal Analysis</td>
<td>Spring</td>
<td>TuTh 2–3:30</td>
</tr>
<tr>
<td>Biostatistics 232 / BST 232: Methods I</td>
<td>Fall</td>
<td>MW 9:45–11:15</td>
</tr>
<tr>
<td>Biostatistics 233 / BST 233: Methods II</td>
<td>Spring</td>
<td>MW 9:45–11:15</td>
</tr>
<tr>
<td>Biostatistics 245 / BST 245(^7): Analysis of Multivariate and Longitudinal Data</td>
<td>Spring</td>
<td>TuTh 9:45–11:15</td>
</tr>
</tbody>
</table>

Last Updated: 2/16/2018
<table>
<thead>
<tr>
<th>Course</th>
<th>Term</th>
<th>Days</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biostatistics 249 / BST 249*: Bayesian Methodology in Biostatistics</td>
<td>Fall</td>
<td>TuTh</td>
<td>9:45-11:15</td>
</tr>
<tr>
<td>BST 263: Applied Machine Learning</td>
<td>Spring</td>
<td>TTh</td>
<td>3:45-5:15</td>
</tr>
<tr>
<td>BIO 515: Measurement Error and Misclassification</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SBS 263: Multilevel Statistical Methods: Concept and Application</td>
<td>Spring</td>
<td>F</td>
<td>9:45-12:45</td>
</tr>
<tr>
<td>ECON 2150: Machine Learning in Econometrics</td>
<td>Spring</td>
<td></td>
<td>TBA</td>
</tr>
<tr>
<td>EDU S-052: Applied Data Analysis</td>
<td>Spring</td>
<td></td>
<td>TuTh 10–11:30</td>
</tr>
<tr>
<td>EDU S-077: Applied Longitudinal Data Analysis</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Calculus/Linear Algebra (Not required)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Math 1b: Calculus, Series, and Differential Equations</td>
<td>Fall/Spring</td>
<td></td>
<td>Many options, see online course catalog</td>
</tr>
<tr>
<td>Mathematics 21a. Multivariable Calculus</td>
<td>Fall/Spring</td>
<td></td>
<td>Many options, see online course catalog</td>
</tr>
<tr>
<td>Mathematics 21b. Linear Algebra and Differential Equations</td>
<td>Fall/Spring</td>
<td></td>
<td>Many options, see online course catalog</td>
</tr>
</tbody>
</table>

*Advanced or very challenging course; please check prerequisites
### EVALUATIVE SCIENCES COURSE REQUIREMENTS

<table>
<thead>
<tr>
<th>Class</th>
<th>Semester</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Research Design and Methods (2.0 credits; both classes required)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health Policy 3080A and 3080B: Graduate Reading Course: Evaluative</td>
<td>Full Year</td>
<td>TBA</td>
</tr>
<tr>
<td>Science and Statistics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GHP 228: Econometric Methods in Impact Evaluation</td>
<td>Spring</td>
<td>F 8–11:15</td>
</tr>
<tr>
<td>MIT 14.387: Advanced Econometrics¹ (alternative to GHP 228)</td>
<td>Spring</td>
<td>TBA</td>
</tr>
<tr>
<td><strong>Specific Methodological Approaches</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>(Take 2.0 credits from courses in any of the eight categories below.)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Survey Research Methods (recommended)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistics 160/260: Design and Analysis of Sample Surveys</td>
<td>Fall</td>
<td>MW 2:30–4</td>
</tr>
<tr>
<td>BST 212: Survey Research Methods in Community Health</td>
<td>Spring</td>
<td>W 3:45–5:15</td>
</tr>
<tr>
<td>Causal Inference (highly recommended)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistics 186: Statistical Methods for Evaluating Causal Effects</td>
<td>NA</td>
<td>Not offered 2017-18</td>
</tr>
<tr>
<td>Statistics 240: Matched Sampling and Study Design</td>
<td>NA</td>
<td>Not offered 2017-18</td>
</tr>
<tr>
<td>EPI 201: Introduction to Epidemiology: Methods 1</td>
<td>Fall 1</td>
<td>TuTh 9:45–11:15</td>
</tr>
<tr>
<td>EPI 289: Models for Causal Inference</td>
<td>Spring 1</td>
<td>MW 9:45–11:15</td>
</tr>
<tr>
<td>Decision Sciences¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RDS 280: Decision Analysis for Health and Medical Practices</td>
<td>Fall 2</td>
<td>TuTh 2–3:30</td>
</tr>
<tr>
<td>RDS 282: Economic Evaluation of Health Policy and Program Management</td>
<td>Spring 2</td>
<td>MW 2–3:30</td>
</tr>
<tr>
<td>RDS 284: Decision Theory</td>
<td>Fall</td>
<td>MW 11:30–1</td>
</tr>
<tr>
<td>RDS 285: Decision Analysis Methods in Public Health and Medicine</td>
<td>Spring 1</td>
<td>MW 2–3:30</td>
</tr>
<tr>
<td>API 302: Analytic Frameworks for Policy</td>
<td>Fall</td>
<td>TuTh 8:45–10; review F at 2:45</td>
</tr>
<tr>
<td>Epidemiology</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EPI 202: Epidemiologic Methods II: Elements of Epidemiologic Research</td>
<td>Fall 2</td>
<td>TuTh 9:45–1</td>
</tr>
<tr>
<td>EPI 203: Study Design in Epidemiologic Research</td>
<td>Spring 2</td>
<td>TuTh 11:30–1</td>
</tr>
<tr>
<td>EPI 204: Analysis of Case-Control, Cohort, and Other Epidemiological Data</td>
<td>Spring 2</td>
<td>TuTh 9:45–11:15</td>
</tr>
<tr>
<td>EPI 207: Advanced Epidemiologic Methods</td>
<td>Fall 1</td>
<td>MW 3:45–5:15</td>
</tr>
<tr>
<td>Program Evaluation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>API 208 / EDU S-020: Program Evaluation: Estimating Program</td>
<td>NA</td>
<td>Not offered 2017-18</td>
</tr>
<tr>
<td>Effectiveness with Empirical Analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>API 211: Program Evaluation</td>
<td>Spring</td>
<td>M 4-7</td>
</tr>
<tr>
<td>HPM 543: Quantitative Methods in Program Evaluation</td>
<td>Spring 2</td>
<td>MW 2–3:30</td>
</tr>
<tr>
<td>Qualitative Research</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GHP 504: Qualitative Research Methods for Global Health</td>
<td>Spring 1</td>
<td>F 9:45-1</td>
</tr>
<tr>
<td>SBS 288: Qualitative Research Methods in Public Health</td>
<td>Fall 1</td>
<td>F 9:45–12:45</td>
</tr>
<tr>
<td>EDU S-504: Introduction to Qualitative Research</td>
<td>Fall</td>
<td>W 1–4</td>
</tr>
<tr>
<td>Network Analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECON 980O: Measuring and Modeling Social Networks</td>
<td>NA</td>
<td>Not offered 2017-18</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistics 140: Design of Experiments</td>
<td>NA</td>
<td>Not offered 2017-18</td>
</tr>
<tr>
<td>Government 2010: Strategies of Political Inquiry</td>
<td>NA</td>
<td>Not offered 2017-18</td>
</tr>
<tr>
<td>SBS 245: Social and Behavioral Research Methods</td>
<td>Fall</td>
<td>TuTh 8:30–9:30</td>
</tr>
<tr>
<td>HBS 4070: Design of Field Research Methods</td>
<td>Spring (or not</td>
<td>TBD</td>
</tr>
<tr>
<td>offered)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

¹ESS students may count decision sciences courses both towards the distribution requirements and towards the Specific Methodological Approaches ESS requirement.
# Suggested Sequence of Statistics Courses

<table>
<thead>
<tr>
<th></th>
<th>Fall 1st year</th>
<th>Spring 1st year</th>
<th>Fall 2nd year</th>
<th>Spring 2nd year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biostatistics (Level 1)</td>
<td>1) BIO 213 or BIO 211</td>
<td>1) BIO 223 or BIO 226</td>
<td>1) BIO 210</td>
<td>1) BIO 223 or BIO 226</td>
</tr>
<tr>
<td></td>
<td>2) BIO 222</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biostatistics (Level 2)</td>
<td>1) BIO 230</td>
<td>1) BIO 231</td>
<td>1) BIO 235</td>
<td>1) BIO 223</td>
</tr>
<tr>
<td></td>
<td>2) BIO 232</td>
<td>2) BIO 233</td>
<td></td>
<td>2) BIO 226</td>
</tr>
<tr>
<td>Statistics (Level 2)</td>
<td>1) STAT 110</td>
<td>1) STAT 111</td>
<td>1) STAT 139</td>
<td>1) STAT 149</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2) GOV 2001</td>
</tr>
<tr>
<td>Statistics (Level 3)</td>
<td>1) STAT 210</td>
<td>1) STAT 211</td>
<td>1) STAT 220</td>
<td></td>
</tr>
<tr>
<td>Econometrics (Level 1)</td>
<td>1) API-209 or GHP 525</td>
<td>1) API-210</td>
<td></td>
<td>1) ECON 2120</td>
</tr>
<tr>
<td>Econometrics (Level 2)</td>
<td>1) ECON 2110</td>
<td>1) ECON 2120</td>
<td>1) ECON 2140</td>
<td></td>
</tr>
</tbody>
</table>
FACULTY MEMBERS ASSOCIATED WITH THE ESS CONCENTRATION

**Concentration Chairs**
Mary Beth Landrum
   Professor of Health Policy (Biostatistics), Harvard Medical School
J. Michael McWilliams
   Warren Alpert Associate Professor of Health Care Policy and Associate Professor of Medicine, Harvard Medical School

**Associated Faculty**
Niteesh Choudhry
   Professor of Medicine, Harvard Medical School
Jessica Cohen
   Associate Professor of Global Health, Harvard T. H. Chan School of Public Health
Benjamin Le Cook
   Assistant Professor of Psychiatry, Harvard Medical School
Arnold M. Epstein
   John H. Foster Professor of Health Policy and Management, Harvard T.H. Chan School of Public Health
   Professor of Medicine and Health Care Policy, Harvard Medical School
Laura Garabedian
   Assistant Professor of Population Medicine, Harvard Medical School and Harvard Pilgrim Healthcare Institute
Laura Hatfield
   Associate Professor of Health Care Policy (Biostatistics), Harvard Medical School
David Hemenway
   Professor of Health Policy, Harvard School of Public Health
John Hsu
   Associate Professor of Medicine, Harvard Medical School
Gary King
   Albert J. Weatherhead III University Professor
Dan Levy
   Senior Lecturer in Public Policy, Harvard Kennedy School
Christine Lu
   Associate Professor of Population Medicine, Harvard Medical School and Harvard Pilgrim Healthcare Institute
Barbara J. McNeil
   Ridley Watts Professor of Health Care Policy, Harvard Medical School
   Professor of Radiology, Harvard Medical School
Carl N. Morris
   Professor of Statistics Emeritus, Faculty of Arts and Sciences
Sharon-Lise Normand
   Professor of Health Care Policy (Biostatistics), Harvard Medical School
   Professor in the Department of Biostatistics, Harvard T.H. Chan School of Public Health
Sherri Rose
   Associate Professor of Health Care Policy (Biostatistics), Harvard Medical School
Dennis Ross-Degnan
   Associate Professor of Population Medicine, Harvard Medical School and Harvard Pilgrim Healthcare Institute
Donald B. Rubin
   John L. Loeb Professor of Statistics, Faculty of Arts and Sciences
Stephen Soumerai
   Professor of Population Medicine, Harvard Medical School and Harvard Pilgrim Healthcare Institute
Theodore Svoronos
   Lecturer in Public Policy, Harvard Kennedy School
Katherine Swartz
   Professor of Health Policy and Economics, Harvard T.H. Chan School of Public Health
James Frank Wharam
   Associate Professor of Population Medicine, Harvard Medical School and Harvard Pilgrim Healthcare Institute
Alan Zaslavsky
   Professor of Health Care Policy (Statistics), Harvard Medical School
<table>
<thead>
<tr>
<th>Course Description</th>
<th>Probability Theory (1.0 credit)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Statistics 110. Introduction to Probability</strong></td>
<td></td>
</tr>
</tbody>
</table>
A comprehensive introduction to probability. Basics: sample spaces and events, conditional probability, and Bayes’ Theorem. 
**Prerequisite:** Mathematics 18 or above (may be taken concurrently).

| **Statistics 171. Introduction to Stochastic Processes** |
A introductory course in stochastic processes. Topics include Markov chains, branching processes, Poisson processes, birth and death processes, Brownian motion, martingales, introduction to stochastic integrals, and their applications. 
**Prerequisite:** Statistics 110 or equivalent.

| **Statistics 210. Probability I** |
**Prerequisite:** Statistics 110 or equivalent required; Statistics 111 or equivalent recommended.

| **API 201 (A,B): Quantitative Analysis and Empirical Methods** |
Introduces students to concepts and techniques essential to the analysis of public policy issues. Provides an introduction to probability, statistics, and decision analysis emphasizing the ways in which these tools are applied to practical policy questions. 
**Prerequisite:** Probability (Stat 110), Linear Algebra and Calculus (Math 21a and 21b)

| **BST 222. Basics of Statistical Inference** |
This course will provide a basic, yet thorough introduction to the probability theory and mathematical statistics that underlie many of the commonly used techniques in public health research. Topics to be covered include probability distributions (normal, binomial, Poisson), means, variances and expected values, finite sampling distributions, parameter estimation (method of moments, maximum likelihood), confidence intervals, hypothesis testing (likelihood ratio, Wald and score tests). All theoretical material will be motivated with problems from epidemiology, biostatistics, environmental health and other public health areas. This course is aimed towards second year doctoral students in fields other than Biostatistics. Background in algebra and calculus required. 
**Course Note:** Lab or section times to be announced at first meeting. 
**Course Prerequisites:** BIO210 or BIO211 or BIO213

| **Biostatistics 230. Probability Theory and Applications I** |
Axiomatic foundations of probability, independence, conditional probability, joint distributions, transformations, moment generating functions, characteristic functions, moment inequalities, sampling distributions, modes of convergence and their interrelationships, laws of large numbers, central limit theorem, and stochastic processes. 
**Note:** Offered jointly with the School of Public Health as BST 230.

| Statistical Inference (1.0 credit) |
Statistics 111. Introduction to Theoretical Statistics
Basic concepts of statistical inference from frequentist and Bayesian perspectives. Topics include maximum likelihood methods, confidence and Bayesian interval estimation, hypothesis testing, least squares methods and categorical data analysis. 
Prerequisite: Mathematics 19a and 19b or equivalent and Statistics 110.

Statistics 186. Statistical Methods for Evaluating Causal Effects
Statistical methods for inferring causal effects from data from randomized experiments or observational studies. Students will develop expertise to assess the credibility of causal claims and the ability to apply the relevant statistical methods for causal analyses. Examples from many disciplines: economics, education, other social sciences, epidemiology, and biomedical science. Evaluations of job training programs, educational voucher schemes, changes in laws such as minimum wage laws, medical treatments, smoking, military service.
Prerequisite: Statistics 100; Mathematics 21a, b or permission of instructor.

Statistics 211. Statistical Inference I
Inference: frequency, Bayes, decision analysis, foundations. Likelihood, sufficiency, and information measures. Models: Normal, exponential families, multilevel, and non-parametric. Point, interval and set estimation; hypothesis tests. Computational strategies, large and moderate sample approximations.
Prerequisite: Statistics 111 and 210a or equivalent.

API 201 (A,B): Quantitative Analysis and Empirical Methods
Introduces students to concepts and techniques essential to the analysis of public policy issues. Provides an introduction to probability, statistics, and decision analysis emphasizing the ways in which these tools are applied to practical policy questions. 
Prerequisite: Probability (Stat 110), Linear Algebra and Calculus (Math 21a and 21b)

Graduate-level version of Gov. 1002. Meets with Gov. 1002, introduces theories of inference underlying most statistical methods and how new approaches are developed. Examples include discrete choice, event counts, durations, missing data, ecological inference, time-series cross sectional analysis, compositional data, causal inference, and others. Will require extra homework and examination problems in addition to those for Gov. 1002.
Prerequisite: Government 2000 or the equivalent.

BST 222. Basics of Statistical Inference
(listed above under Probability Theory)

Biostatistics 231. Statistical Inference I
Exponential families, sufficiency, ancillarity, completeness, method of moments, maximum likelihood, unbiased estimation, Rao-Blackwell and Lehmann-Scheffe theorems, information inequality, Neyman-Pearson theory, likelihood ratio, score and Wald tests, uniformly and locally most powerful tests, asymptotic relative efficiency.
Note: Offered jointly with the School of Public Health as BST 231.
Prerequisite: Biostatistics 230 or signature of instructor required.

API 209. Advanced Quantitative Methods I: Statistics
The goal of this course is to prepare students to analyze public policy issues using statistics. Topics included fall in the areas of probability theory, sampling, estimation, hypothesis testing, and regression analysis. While many students taking this class will have already taken courses in statistics and regression analysis, this course will probably place a much stronger emphasis than typical courses on conceptually understanding the statistical methods. Since the course is targeted to first-year students in the MPA/ID program, we will not shy away from using the mathematical tools needed to develop the conceptual understanding. But the emphasis of the course will be on the conceptual understanding and application of the tools rather than on the math or the mechanics behind the tools.
Note: This course is open to non-MPA/ID students only by permission of the instructor. May not be taken for credit with API-201.
Prerequisite: Multivariate calculus or linear algebra.
Regression (1.0 credit)

Statistics 139. Statistical Sleuthing Through Linear Models
A serious introduction to statistical inference with linear models and related methods. Topics include t-tools and permutation-based alternatives, multiple-group comparisons, analysis of variance, linear regression, model checking and refinement, and causation versus correlation. Emphasis on thinking statistically, evaluating assumptions, and developing tools for real-life applications.
Prerequisite: Statistics 100 or equivalent and Mathematics 19a and 19b or equivalent.

Statistics 244. Linear and Generalized Linear Models
The theory and application of linear and generalized linear models, including linear models for normal responses, logistic models for binary and multinomial data, loglinear models for count data, overdispersion and quasi likelihood methods, and models and methods for clustered (e.g., repeated measurement) correlated data.
Prerequisite: Strong statistics background required (at the second-year graduate level), Statistics 210a may be taken concurrently, Statistics 211a desirable.

Economics 1123. Introduction to Econometrics
An introduction to multiple regression techniques with focus on economic applications. Discusses extensions to discrete response, panel data, and time series models, as well as issues such as omitted variables, missing data.
Prerequisite: Statistics 100.

API 201 (C, D): Quantitative Analysis and Empirical Methods
Introduces students to concepts and techniques essential to the analysis of public policy issues. Provides an introduction to probability, statistics, and decision analysis emphasizing the ways in which these tools are applied to practical policy questions.
Prerequisite: API 201 (A,B)

Government 2000. Introduction to Quantitative Methods I
Graduate-level version of Government 1000. Meets with Government 1000, an introduction to statistical research in political science with a focus on applied linear regression. Will require extra homework and examination problems in addition to those for Government 1000.
Prerequisite: Permission of the instructor for anyone other than Government Department graduate students.

BST 211. Regression and Analysis of Variance in Experimental Research
Covers analysis of variance and regression, including details of data-analytic techniques and implications for study design. Also included are probability models and computing. Students learn to formulate a scientific question in terms of a statistical model, leading to objective and quantitative answers.
Course Note: Lab or section times to be announced at first meeting.
Course Prerequisites: BIO200 or BIO201 or ID200 or ID207 or BIO202&203 or BIO206&207 or BIO206&208 or BIO206&209

BST 213. Applied Regression for Clinical Research
This course will introduce students involved with clinical research to the practical application of multiple regression analysis. Linear regression, logistic regression and proportional hazards survival models will be covered, as well as general concepts in model selection, goodness-of-fit, and testing procedures. Each lecture will be accompanied by a data analysis using SAS and a classroom discussion of the results. The course will introduce, but will not attempt to develop the underlying likelihood theory. Background in SAS programming ability required.
Note: Lab or section times to be announced at first meeting.
Prerequisites: BIO200 or BIO201 or ID200 or ID207 or BIO202&203 or BIO206&207 or BIO206&208 or BIO206&209

Biostatistics 232. Methods I
Introductory course in the analysis of Gaussian and categorical data. The general linear regression model, ANOVA, robust alternatives based on permutations, model building, resampling methods (bootstrap and jackknife), contingency tables, exact methods, logistic regression.
Note: Offered jointly with the School of Public Health as BST 232.
Prerequisite: Signature of instructor required.

*Biostatistics 235. Advanced Regression and Statistical Learning
An advanced course in linear models, including both classical theory and methods for high dimensional data. Topics include theory of estimation and hypothesis testing, multiple testing problems and false discovery rates, cross validation and model selection, regularization and the LASSO, principal components and dimension reduction, and classification methods. Background in matrix algebra and linear regression required.
GHP 525. Econometrics for Health Policy
This is a course in applied econometrics for doctoral and advanced master level students. The course has two primary objectives: (1) to develop skills in linking economic behavioral models and quantitative analysis, in a way that students can use in their own research; (2) to develop students' abilities to understand and evaluate critically other peoples' econometric studies. The course focuses on developing the theoretical basis and practical application of the most common empirical models used in health policy research. In particular, it pays special attention to a class of models identifying causal effects in observational data, including instrumental variable estimation, simultaneous equations and two-stage-least-squares, quasi-experiments and difference-in-difference method, sample selection, treatment effect models and propensity score methods. Lectures will be complemented with computer exercises building on public domain data sets commonly used in health research. The statistical package recommended for the exercises is Stata.

Course Note: Students are expected to be familiar with probability theory (density and distribution functions) as well as the concepts underlying basic ordinary least square (OLS) estimation.

Course Activities: Optional review and computer lab sessions will be held.

Course Prerequisites: BIO210 or BIO211 or BIO213

API 210. Advanced Quantitative Methods II: Econometric Methods
Intended as a continuation of API-209, Advanced Quantitative Methods I, this course focuses on developing the theoretical basis and practical application of the most common tools of empirical analysis including non-linear models, instrumental variables, and panel data. Foundations of analysis will be coupled with hands-on examples and assignments involving the analysis of data sets.

Note: This course is open to non-MPA/ID students only by permission of the instructor. May not be taken for credit with API-202.

Prerequisite: API-209 or permission of instructor.

Other Data Analysis (2.0 credits)

Statistics 131. Time Series and Prediction
Introduction to time series models and associated methods of data analysis and inference. Auto regressive (AR), moving average (MA), ARMA, and ARIMA processes, stationary and non-stationary processes, seasonal processes, auto-correlation and partial auto-correlation functions, identification of models, estimation of parameters, diagnostic checking of fitted models, forecasting, spectral analysis, and transfer function models.

Prerequisite: Statistics 111 and 139 or equivalent.

Statistics 149. Statistical Sleuthing through Generalized Linear Models
Sequel to Statistics 139, emphasizing common methods for analyzing continuous non-normal and categorical data. Topics include logistic regression, log-linear models, multinomial logit models, proportional odds models for ordinal data, Gamma and inverse-Gaussian models, over-dispersion, analysis of deviance, model selection and criticism, model diagnostics, and an introduction to non-parametric regression methods.

Note: Examples will be drawn from several fields, particularly from biology and social sciences.

Prerequisite: Statistics 139 or with permission of instructor.

Statistics 151. Multilevel or Longitudinal Models
Data often have structure that needs to be modeled explicitly. For example, when investigating students' outcomes we need to account for the fact that students are nested inside classes that are in turn nested inside schools. If we are watching students develop over time, we need to account for the dependence of measurements across time. If we do not, our inferences will tend to be overly optimistic and wrong. The course provides an overall framework, the multilevel and generalized multilevel (hierarchical) model, for thinking about and analyzing these forms of data. We will focus on specific versions of these tools for the most common forms of longitudinal and clustered data. This course will focus on applied work, using real data sets and the statistical software R. R will be specifically taught and supported. While the primary focus will be on the linear model with continuous outcomes (i.e., the classic regression framework) we will also discuss binary, categorical, and ordinal outcomes. We will emphasize how to think about the applicability of these methods, how they might fail, and what one might do to protect oneself in such circumstances. Applications of hierarchical (multi-level) models will include the canonical specific cases of random-slope, random-intercept, mixed effect, crossed effect, marginal, and growth-curve models.

Permission of instructor required.

Prerequisite: S-052, Stat 139, or an equivalent. Jointly-offered in the Graduate School of Education.

Statistics 160/260. Design and Analysis of Sample Surveys
Methods for design and analysis of sample surveys. The toolkit of sample design features and their use in optimal design strategies. Sampling weights and variance estimation methods, including resampling methods. Brief overview of nonstatistical aspects of survey methodology such as survey administration and questionnaire design and validation (quantitative and qualitative). Additional topics: calibration estimators, variance estimation for complex surveys and estimators, nonresponse, missing data, hierarchical models, and small-area estimation.

Prerequisite: Statistics 111 or 139 or with permission of instructor.

Statistics 220. Bayesian Data Analysis
Basic Bayesian models, followed by more complicated hierarchical and mixture models with nonstandard solutions. Includes methods for monitoring adequacy of models and examining sensitivity of models.

Note: Emphasis throughout term on drawing inferences via computer simulation rather than mathematical analysis.

Prerequisite: Statistics 110 and 111.

Statistics 232R. Topics in Missing Data
The modern era of work on missing data problems began in the 1970s and has seen an explosion of developments since then. Seminar will focus on an updated version of a classic text, supplemented with classic articles.

Economics 2140. Econometric Methods
Econometric methods for cross-section and panel data. Topics include generalized method of moments, empirical likelihood, instrumental variables, bootstrapping, clustering, treatment effects, selection bias, difference-in-differences, qualitative choice, quantile regression, nonparametric methods, and semiparametric methods.

Prerequisite: Economics 2120 or equivalent.

(listed above under Statistical Inference)

Government 2735. Empirical Models in International Relations
This course examines statistical issues relevant to the study of international politics. The purpose is to familiarize students with different models that have been employed in research on international conflict, IPE and international institutions.

**Sociology 211 (formerly Sociology 203a), Analysis of Longitudinal Data: Seminar**
This course takes an applied approach to the analysis of longitudinal data. Lectures will provide an overview of a variety of techniques, including fixed effects models, multilevel models, and duration models. Students will develop their own empirical projects and receive support as they begin to work with longitudinal datasets

*Note:* Primarily for graduate students in sociology.

**BST 210. Applied Regression**
Emphasizes concepts and methods for analysis of data which are categorical, rate-of-occurrence (e.g., incidence rate), and time-to-event (survival duration). Stresses applications in epidemiology, clinical trials, and other public health research. Topics include measures of association, 2x2 tables, stratification, matched pairs, logistic regression, model building, analysis of rates, and survival data analysis using proportional hazards models.

*Course Note:* Lab or section times to be announced at first meeting.

*Course Prerequisite(s):* BIO200 or BIO201 or ID200 or ID207 or BIO202&203 or BIO206&207 or BIO206&208 or BIO206&209

**BST 223. Applied Survival Analysis**
Topics will include types of censoring, hazard, survivor, and cumulative hazard functions, Kaplan-Meier and actuarial estimation of the survival distribution, comparison of survival using log rank and other tests, regression models including the Cox proportional hazards model and the accelerated failure time model, adjustment for time-varying covariates, and the use of parametric distributions (exponential, Weibull) in survival analysis. Methods for recurrent survival outcomes and competing risks will also be discussed, as well as design of studies with survival outcomes. Class material will include presentation of statistical methods for estimation and testing along with current software (SAS, Stata) for implementing analyses of survival data. Applications to real data will be emphasized.

*Course Prerequisite(s):* BIO210 or BIO213 or BIO230

**BST 226. Applied Longitudinal Analysis**
This course covers modern methods for the analysis of repeated measures, correlated outcomes and longitudinal data, including the unbalanced and incomplete data sets characteristic of biomedical research. Topics include an introduction to the analysis of correlated data, analysis of response profiles, fitting parametric curves, covariance pattern models, random effects and growth curve models, and generalized linear models for longitudinal data, including generalized estimating equations (GEE) and generalized linear mixed effects models (GLMMs).

*Course Activities:* Homework assignments will focus on data analysis in SAS using PROC GLM, PROC MIXED, PROC GENMOD, and PROC NLMIXED.

*Course Note:* Lab or section times will be announced at first meeting.

*Course Prerequisite(s):* BIO210 or BIO211 or BIO213 or BIO232

**Biostatistics 232. Methods I**
(listed above under Regression)

**Biostatistics 233. Methods II**
Intermediate course in the analysis of Gaussian, categorical, and survival data. The generalized linear model, Poisson regression, random effects and mixed models, comparing survival distributions, proportional hazards regression, splines and smoothing, the generalized additive model.

*Note:* Offered jointly with the School of Public Health as BST 233.

*Prerequisite:* Biostatistics 232 or signature of instructor required.

**Biostatistics 245. Analysis of Multivariate and Longitudinal Data**
The multivariate normal distribution, Hotelling’s T2, MANOVA, repeated measures, the multivariate linear model, random effects and growth curve models, generalized estimating equations, multivariate categorical outcomes, missing data, computational issues for traditional and new methodologies.

*Note:* Offered jointly with the School of Public Health as BST 245.

*Prerequisite:* Biostatistics 231 and Biostatistics 235.

**Biostatistics 249. Bayesian Methodology in Biostatistics**
General principles of the Bayesian approach, prior distributions, hierarchical models and modeling techniques, approximate inference, Markov chain Monte Carlo methods, model assessment and comparison. Bayesian approaches to GLMMs, multiple testing, nonparametrics, clinical trials, survival analysis.

*Note: Offered jointly with the School of Public Health as BST 249.*

**Prerequisite:** Biostatistics 231 and Biostatistics 232, or signature of instructor required.

**BST 263: Applied Machine Learning**

The central theme of the course will be to ground the material in practical real-world data examples, in order to motivate the concepts and illustrate why and how the methods work. Some mathematical foundations will be covered, but the primary emphasis of the course will be on learning how to implement and use the methods, while gaining an intuitive understanding of them. Programming (in R) and case studies will be used throughout the course to provide hands-on training.

**BST 515. Measurement Error and Misclassification**

This course will cover theory for valid estimation and inference in statistical analysis when covariates are mis-classified or measured with error. Methods for contingency tables, generalized linear models and for survival analysis will be addressed. Topics include likelihood-based methods, regression calibration, SIMEX, instrumental variables and optimal study design.

**Course Prerequisites:** BIO231 and BIO223 required

**Course Note:** Course is mutually exclusive with EPI515. You may not take both this course and EPI515.

**SBS 263. Multilevel Statistical Methods: Concept and Application**

This course is designed to provide doctoral students with a training experience in the concept and application of multilevel statistical modeling. Students will be motivated to think about correlated and dependent data structures that arise due to sampling design and/or are inherent in the population (such as pupils nested within schools; patients nested within clinics; individuals nested within neighborhoods and so on). The substantive motivation for analyzing such complex data structures would be to make quantitative assessments about the role of contexts (e.g., schools, clinics, neighborhoods) in predicting individual outcomes. In particular, the principles of recognizing and modeling the underlying heterogeneity in average relationships would be emphasized. Linear, non-linear, and multivariate multilevel models will be covered. Upon completion, students should be able to conceptualize multilevel modeling strategies and to undertake empirical, quantitative multilevel research. The course will be lecture-based with substantial hands-on component.

**Course Activities:** Data management, modeling and analysis; individual assignments; project submission and class participation.

**Course Prerequisite(s):** SBS245

**Course Notes:** This course is a requirement for all SBS doctoral students. Required lab.

**Econ 2150. Machine Learning in Econometrics: Prediction, Estimation, and Big Data**

Innovations in machine learning ('big data') have created many engineering breakthroughs from real time voice recognition to automatic categorization (and in some cases production) of news stories. Since these techniques are at their essence novel ways to work with data, they should also have implications for social science. This course explores the intersection of machine learning and social science and aims to answer a few questions about these new techniques: (i) How do they work and what kinds of statistical guarantees can be made about their performance? (ii) How can they be used to answer questions that interest social science researchers, such as testing theories or improving social policy; and (iii) How might they open up new research questions? We will cover standard machine learning techniques such as supervised and unsupervised learning, statistical learning theory and nonparametric and Bayesian approaches. The goal is to create a working understanding of when and how they can be profitably applied. Students will be required to apply some of these techniques themselves, but we will not cover the computational aspects of the underlying methods. The course is aimed at PhD students with a solid background in statistical techniques, such as comes from the equivalent of a first year economics PhD econometrics sequence.

**EDU S-052. Applied Data Analysis**

This course is designed for those who want to extend their data analytic skills beyond a basic knowledge of multiple regression analysis and who want to communicate their findings clearly to audiences of researchers, scholars, and policymakers. The course contributes directly to the diverse data analytic toolkit that the well-equipped empirical researcher must possess in order to perform sensible analyses of complex educational, psychological, and social data. Topics in the course include more extensive use of transformations in regression analysis, influence statistics, building and comparing taxonomies of regression models, general linear hypothesis testing, an introduction to multilevel modeling, nonlinear regression analysis, binomial logistic regression analysis, principal components analysis, cluster analysis, an introduction to discrete time survival analysis, and others. S-052 is an applied course that offers conceptual explanations of statistical techniques, along with opportunities to examine, implement, and practice them in real data. Because the course will feature the intensive use of Stata statistical software in all data analyses, learning the computer skills necessary to conduct these kinds of analyses, and the communication skills to discuss them, is an integral part of the course. Weekly section attendance is strongly encouraged.

**Prerequisite:** Successful completion of S-040 or an equivalent course covering applied multiple regression.
EDU S-077. Applied Longitudinal Data Analysis

Researchers in education and the social sciences often pose research questions about change and event occurrence over time. For instance, a researcher investigating the development of reading skills in young children might ask: How rapidly do children’s reading skills develop as they age, and do the skills of boys and girls develop at different rates? Alternatively, her questions could be framed in terms of whether and when children achieve particular developmental milestones. Then, she would ask: When does a child make the transition from "learning to read" to "reading to learn," and do children exposed to innovative reading programs make the transition at different ages? Answering these kinds of questions requires longitudinal (panel) data and the application of innovative statistical methods. The goal of S-077 is to introduce two of these methods: individual growth modeling and survival analysis. The course is a seminar in which class members must take individual responsibility for reading new material, preparing written answers to discussion questions, and providing feedback to peers. Class members will also conduct an original research project of their own devising, using the new methods; make a public presentation of their findings; and submit a final paper of their original research. The nature of the research projects is negotiable, but must directly advance class members’ professional and scholarly agenda. Enrollment is limited to advanced students who have completed the S-052 course, have arranged access to appropriate longitudinal data, and have proposed a suitable research project.
Course Descriptions: Evaluative Sciences

**Research Design and Methodology (2.0 credits; both classes required)**

**Health Policy 3080A&B. Graduate Reading Course: Evaluative Science and Statistics**
This course will include readings on study designs that help develop critical analysis skills. It will also include a close reading of the Shadish, Cook, and Campbell book. During the second semester, the course will prepare students for the ESS qualifying exam.

**GHP 228. Econometric Methods in Impact Evaluation**
The objective of this course is to provide students with a set of theoretical, econometric and reasoning skills to estimate the causal impact of one variable on another. Examples from the readings explore the causal effect of policies, laws, programs and natural experiments derived from pension programs to television shows to natural disasters. We will go beyond estimating causal effects to analyze the channels through which the causal impact was likely achieved. This will require that the students are familiar with microeconomic theories of incentives, institutions, social networks, etc.
The course will introduce students to a variety of econometric techniques in impact evaluation and a set of reasoning skills intended to help them become both a consumer and producer of applied empirical research. Students will learn to critically analyze evaluation research and to gauge how convincing the research is in identifying a causal impact. They will use these skills to develop an evaluation plan for a topic of their own, with the aim of stimulating ideas for dissertation research. This is a methods class that relies heavily on familiarity with econometrics and microeconomics. These are pre-requisites for the course without exception. The course is intended for doctoral students who are finishing their course work and aims to help them transition into independent research.
The aim of this course is to prepare doctoral students in the health systems track of the Global Health and Population department for the dissertation phase of their research and thus they will be given priority in enrollment. The course is also open to other GHP doctoral students, other GHP masters students and students from other departments, conditional on having adequate training in economics and the course having enough space.
Prerequisite: Econometrics and intermediate micro-economics (GHP 525 and GHP 291 or equivalent) are required for this course. While students can get by with just these two subjects, some previous experience with regression analysis and applied economic research will be a huge advantage. Students seeing applied regression analysis for the first time in this course will most likely struggle with the reading.

**Specific Methodological Approaches (2.0 credits chosen from Survey Research Methods, Causal Inference, Decision Sciences, Epidemiology, Program Evaluation, Qualitative Research, Network Analysis, and Other)**

**Survey Research Methods**

**Statistics 160. Design and Analysis of Sample Surveys**
(listed above under Other Data Analyses)

**BST 212. Survey Research Methods In Community Health**
Covers research design, sample selection, questionnaire construction, interviewing techniques, the reduction and interpretation of data, and related facets of population survey investigations. Focuses primarily on the application of survey methods to problems of health program planning and evaluation. Treatment of methodology is sufficiently broad to be suitable for students who are concerned with epidemiological, nutritional, or other types of survey research.
Causal Inference

Statistics 186. Statistical Methods for Evaluating Causal Effects
(listed above under Statistical Inference)

Statistics 240. Matched Sampling and Study Design
This course provides an accessible introduction to the study of matched sampling and other design techniques in any field (e.g., economics, education, epidemiology, medicine, political science, etc.) conducting empirical research to evaluate the causal effects of interventions.

Prerequisite: Statistics 110, Statistics 111, and Statistics 139.

EPI 201. Epidemiologic Methods I
EPI201 introduces the principles and methods used in epidemiologic research. The course discusses the conceptual and practical issues encountered in the design and analysis of epidemiologic studies for description and causal inference. The final exam requires the application of the learned skills to a real problem in epidemiology. EPI201 is the first course in the series of methods courses designed for students majoring in Epidemiology or Biostatistics, and those interested in a detailed introduction to the design and conduct of epidemiologic studies. Students who take EPI201 are expected to take EPI202 (Methods II).

Course Note: Thursday or Friday lab required.

Course is mutually exclusive with EPI200, EPI208, EPI500, EPI505, ID200, and ID538. You may not take both this course and any of those courses.

EPI 289. Models for Causal Inference
EPI289 describes models for causal inference, their assumptions, and their practical application to epidemiologic data. The course covers propensity score methods, the parametric g-formula, inverse probability weighting of marginal structural models, g-estimation of nested structural models, and instrumental variable methods. The course also introduces models for causal inference in the presence of time-varying exposures, which will be extensively studied in EPI207. EPI289 is designed to be taken after EPI201/EPI202. The epidemiologic concepts and methods studied in EPI201/202 will be reformulated within a modeling framework in EPI289. Familiarity with the SAS language is strongly recommended.

Course Prerequisite(s): EPI201 and EPI202

Decision Sciences

RDS 280. Decision Analysis for Health and Medical Practices
This course is designed to introduce the student to the methods and growing range of applications of decision analysis and cost-effectiveness analysis in health technology assessment, medical and public health decision making, and health resource allocation. The objectives of the course are: (1) to provide a basic technical understanding of the methods used, (2) to give the student an appreciation of the practical problems in applying these methods to the evaluation of clinical interventions and public health policies, and (3) to give the student an appreciation of the uses and limitations of these methods in decision making at the individual, organizational, and policy level both in developed and developing countries.

Course Note: Introductory economics is recommended but not required.

Course Prerequisites: ID538 or BIO200 or BIO201 or BIO202&203 or BIO206&207 or BIO206&208 or BIO206&209 (all courses may be taken concurrently)

RDS 282. Economic Evaluation of Health Policy and Program Management
This course features case studies in the application of health decision science to policymaking and program management at various levels of the health system. Both developed and developing country contexts will be covered. Topics include: [1] theoretical foundations of cost-effectiveness analysis (CEA); [2] controversies and limitations of CEA in practice; [3] design and implementation of tools and protocols for measurement and valuation of cost and benefit of health programs; [4] integration of evidence of economic value into strategic planning and resource allocation decisions, performance monitoring and program evaluation; [5] the role of evidence of economic value in the context of other stakeholder criteria and political motivations.

Course Prerequisites: Students must have taken RDS280 or RDS286. Prior coursework in Microeconomics is recommended.

RDS 284. Decision Theory
Introduces the standard model of decision-making under uncertainty, its conceptual foundations, challenges, alternatives, and methodological issues arising from the application of these techniques to health issues. Topics include von Neumann-Morgenstern and multi-attribute utility theory, Bayesian statistical decision theory, stochastic dominance, the value of information, judgment under uncertainty and alternative models of probability and decision making (regret theory, prospect
theory, generalized expected utility). Applications are to preferences for health and aggregation of preferences over time and across individuals.

RDS 285. Decision Analysis Methods in Public Health and Medicine
An intermediate-level course on methods and health applications of decision analysis modeling techniques. Topics include Markov models, microsimulation models, life expectancy estimation, deterministic and probabilistic sensitivity analysis, ROC analysis and diagnostic technology assessment, and cost-effectiveness analysis. 
Course Note: Familiarity with matrix algebra and elementary calculus may be helpful but not required; lab or section times to be announced at first meeting. 
Course Prerequisites: (BIO200 or ID200 or BIO201) and (RDS280 or RDS286)

API-302. Analytic Frameworks for Policy
This course develops abilities in using analytic frameworks in the formulation and assessment of public policies. It considers a variety of analytic techniques, particularly those directed toward uncertainty and interactive decision problems. It emphasizes the application of techniques to policy analysis, not formal derivations. Students encounter case studies, methodological readings, modeling of current events, the computer, a final exam, and challenging problem sets.
Prerequisites: An understanding of intermediate-level microeconomic theory and introductory techniques of optimization and decision analysis; API-101, API-102, or equivalent. Also offered by the Department of Economics as Ec 1415.

Epidemiology

EPI 202. Epidemiologic Methods II: Elements of Epidemiologic Research
Introduces elements of study design, data analysis and inference in epidemiologic research. Principles and methods are illustrated with examples, and reviewed through homework and in-class exercises. May serve as an introduction to more advanced study or as a concluding course for those desiring a working knowledge of epidemiologic methods. EPI 202 extends the concepts of study design, data analysis, and inference introduced in EPI201.
Course Prerequisites: (EPI201 or EPI208 or EPI500 or ID200 or ID207) and (BIO200 or BIO201 or ID200 or ID207 or BIO202&203 or BIO206&207/8/9) (all courses may be taken concurrently)

EPI 203. Study Design in Epidemiologic Research
Beginning with the randomized clinical trial as a paradigm, this course examines common problems in the design, analysis, and interpretation of observational studies. Cohort and case-control studies are the focus of the discussion, but not to the exclusion of other designs. Problems of exposure and disease definitions, time-dependent effects, confounding, and misclassification are considered in the light of data sources typically available. Relevant statistical methods are introduced but not developed in detail.
Course Prerequisites: EPI202 and (BIO200 or ID200 or BIO201 or BIO202&203 or BIO206&207/8/9)

EPI 204. Analysis of Case Control, Cohort, and Other Epidemiologic Data
This course will examine, through practical examples, the use of regression methods for analyses of epidemiologic data, primarily case-control and cohort studies. Methods used will include linear, logistic, Poisson, conditional logistic and Cox regression models. The lectures will focus on the principle ideas and issues underlying the regression analyses, and the computer labs will provide practical experience applying those methods, using SAS software. Issues to be dealt with include dose-response, confounding, influence, and interaction. It will emphasize analysis and interpretation of results in the context of the study design. Familiarity with basic SAS is required, as this will be used in the labs. This can be met through BIO 113 (Introduction to Data Management and Programming in SAS) or other significant SAS experience.
Course Activities: Written group projects, class discussion, quizzes, homework.
Course Note: Computer lab is required, please sign up for one lab session when registering.
Course Prerequisites: (BIO210 (concurrent enrollment allowed) or BIO213) and (EPI202) and (EPI200 or EPI201 or EPI208 or EPI50)

EPI 207. Advanced Epidemiologic Methods
Provides an in-depth investigation of statistical methods for drawing causal inferences from observational studies. Informal epidemiologic concepts such as confounding, selection bias, overall effects, direct effects, and intermediate variables will be formally defined within the context of a counterfactual causal model and with the help of causal diagrams. Methods for the analysis of the causal effects of time-varying exposures in the presence of time dependent covariates that are simultaneously confounders and intermediate variables will be emphasized. These methods include g-computation algorithm estimators, inverse probability weighted estimators of marginal structural models, g-estimation of structural nested models. As a practicum, students will reanalyze data sets using the above methods.
Course Activities: Class discussion, homework, practicum and final examination.
Course Note: Familiarity with logistic regression and survival analysis is expected; lab time will be announced at first meeting.
Course Prerequisites: EPI204 or (BST 210 and EPI289) or BIO233
Program Evaluation

Program evaluation comprises a set of statistical tools for assessing the impact of public interventions. This methodological course will develop students’ skills in quantitative program evaluation. Students will study a variety of evaluation designs (from random assignment to quasi-experimental evaluation methods) and analyze data from actual evaluations, such as the national Job Training Partnership Act Study. The course evaluates the strengths and weaknesses of alternative evaluation methods. This course meets the PhD requirement for empirical methods.

Prerequisite: Familiarity with the basic concepts of statistical inference and regression analysis (such as API-202 or API-210).

API-211. Program Evaluation
As school districts and state agencies accumulate quantitative student outcome data, demand for evidence of impact will grow. All of us must learn to be critical consumers of quantitative evidence of impact. The key challenge when evaluating the impact of an education policy or program is to identify what would have happened if that policy or program had not been implemented. There are a number of different approaches to constructing a plausible estimate of what would have happened, using experimental or quasi-experimental techniques. In this course, we will have three goals: to gain insight into the strengths and weaknesses of different evaluation designs, including experimental and quasi-experimental techniques; to develop the skills required to be a critical reader of impact evaluations; and to develop the ability to more clearly recognize opportunities for impact evaluations in education and to implement policies in a manner that would be amenable to evaluation. During the course, we will read and critique a number of impact evaluations, replicate the results of several evaluations, and design evaluations of educational programs. The course will focus on quantitative impact evaluations, as opposed to qualitative or process evaluations.

Prerequisite: Successful completion of S-030 or S-040, or prior equivalent training in multiple regression. Also offered by the Graduate School of Education as A-164. Permission of the instructor required. Enrollment procedure will be posted on the HGSE course website.

HPM 543. Quantitative Methods in Program Evaluation
This course will give students the tools that they need to evaluate policy interventions, social programs, and health initiatives. Did the program achieve its goals? Did it reach its target audience? Could it have been more effective? In order to answer these questions, students will develop a flexible set of analytical tools, including both the ability to design an evaluation study and the ability to evaluate existing studies critically.
By the end of the course students will be able to construct a well-designed study to answer well-posed questions, gauge the adequacy of available data, implement an econometric analysis, interpret the results of such studies, and draw policy implications. The course will focus on health policies and programs such as public insurance expansions and public health campaigns, but the techniques will be broadly applicable to other realms such as welfare or education.

Course Note: The material in this course is inherently quantitative, and builds on a base of statistics fundamentals. The prerequisite is a course in basic statistics and probability, such as BIO 200, BIO 201, BIO 202/203, ID 538, ID201 or equivalent. This includes knowledge of confidence intervals and hypothesis testing. It also includes familiarity with the statistical package of your choice- ideally STATA, but SAS or SPSS are fine. During the course students will be given data sets to analyze, but there will be no instruction on the mechanics of opening and manipulating the data with a statistical software package. Students should contact instructor if they are uncertain about whether they have adequate preparation for the class.

Prerequisites: BIO200 or BIO201 or BIO202&203, or BIO206 & (207 or 208 or 209) or ID538 or ID201 or equivalent

Qualitative Research

GHP 504. Qualitative Methods for Global Health
The aim of this course is to provide students with an introduction to qualitative methods for global health research. The module is designed to expose students to a wide range of topics including: developing research questions, sampling and site selection, frequently used qualitative methods (such as interviews, observations, focus groups), design of qualitative research protocols, as well as data management and analysis. Students will engage in a variety of active learning exercises (such as constructing and conducting a short informal interview) and will work in small groups on the preparation of a qualitative research project on a defined topic area of international or multicultural health. Class activities and discussions will aim at building a research community in the class, where students support each other’s development as researchers recognizing the complexity, benefits and limitations of conducting cross-cultural qualitative research.

Course Prerequisite: Prospective students wishing to enroll in GHP 504 must email an essay (maxim half-page) to course TAs, by a specified due date usually in November of the fall semester (check HSPH website for update). All admitted students will be notified before winter break in December. The essay should describe: -Current departmental affiliation, degree program and
remaining time to graduation; Rationale for and interest in pursuing training in qualitative methods; Upcoming plans to use qualitative methods in research; Any prior training in or experience with using qualitative methods in field research (and lessons learned if relevant);
Research topics and populations in which the student plans to use qualitative methods

**SBS 288. Qualitative Research Methods in Public Health**
Qualitative research can be used alone or in combination with quantitative research to investigate public health questions. This introductory-level course begins by examining the variety of potential uses of qualitative methods in public health research and diverse qualitative research approaches. The course then explores specific topics, including: "entering" the community to conduct qualitative research; applying theory to study design and open-ended questions; ensuring study rigor; developing theory-based research questions, specific data collection methods (including, but not limited to, semi-structured interviews, focus groups, participant observation); sampling for qualitative studies; data management; data analysis; writing results and research proposals; and considerations for choosing qualitative methods at each stage of a mixed-methods qualitative or mixed-methods qualitative/quantitative study. Students will be required to participate in class discussions, apply concepts covered in class through assignments to collect and analyze qualitative data, critique qualitative works, and propose a qualitative study.

**EDU S-504. Introduction to Qualitative Research**
This introductory methods course offers students a sense of the terrain of qualitative research, including some of the different tools and approaches available to researchers in the field of education. The assigned readings will include scholarship on the practice and philosophical underpinnings of qualitative research, varied examples of published qualitative research, and raw data. Class sessions will generally follow a workshop format with discussions and activities related to weekly readings. In addition, students will get a feel for the overall process of conducting qualitative research by developing an original research proposal that is informed by preliminary data gathering and analysis. Students will start to develop skills related to designing a study, collecting and analyzing data, making appropriate claims, positioning their work relative to existing literature, and appraising others' qualitative research. Students will also begin to think about their own identities and ethical responsibilities as educational researchers, and develop skills for further and ongoing reflection about their work and their relationship to it.
*Note:* Permission of instructor required. Required for first-year Ph.D. students. Other doctoral students may enroll. A limited number of Ed.M. and C.A.S. students may enroll with permission; the enrollment procedure will be posted on the course website.

**Network Analysis**

**API-309. Networks, Complexity and Their Applications**
The course will cover Network Science, Complexity Science and their applications to the understanding of social, natural and economic systems. The course will equip students with an understanding of conceptual and methodological tools for the empirical analysis of large systems, including, but not limited to, Network Analysis, Chaos Theory and Fractals, Evolutionary Theory and Self Organized Criticality. All of these topics will be covered by combining regular lectures and discussion sessions. Occasionally, the course will be complemented by outside speakers. The course will provide students new analytical tools and perspectives that can be used in problem solving situations. In particular the course will discuss in detail applications of Network Science to the understanding of country's exports structures, the identification of potential priority sectors, industrial policy and structural transformations. Also offered by MIT as MAS 961 and held at MIT.

**Statistics 140. Design of Experiments**
Statistical designs for efficient experimentation in the physical, life, social and management sciences and in engineering. A systematic approach to explore input-output relationships by deliberately manipulating input variables. Topics include completely randomized and randomized block designs, Latin square designs, balanced incomplete block designs, factorial designs, confounding in blocks, fractional replications, and re-randomization. Each topic motivated by real-life examples.
*Prerequisite:* Statistics 111 or 139.

**Government 2010. Strategies for Political Inquiry**
Research design for causal inference in qualitative and quantitative studies. Topics covered include measurement, conceptualization, case studies, the relationship between large-n and small-n studies, process-tracing, surveys, field experiments, and natural experiments, with examples of their use in political science.
*Note:* Primarily for graduate students; may also be taken by undergraduates preparing for senior thesis research.
**SBS 245. Social and Behavioral Research Methods**

Provides a broad overview of social and behavioral research methodology, including experimental, quasi-experimental and non-experimental research design, measurement, sampling, data collection, and testing causal theories. By case studies, methodological readings, discussion, written assignments, and data analytic homeworks students learn to conduct social and behavioral research and more applied program evaluations. Homework includes analytic work with observational and experimental studies and development of new measures.

*Course Activities:* Assigned readings, class participation, homeworks, reflections, two papers.

*Course Note:* a multivariate statistics course strongly recommended; course primarily for doctoral students.

*Course Prerequisites:* BIO210 or BIO211 or BIO213

---

**HBS 4070. Design of Field Research Methods**

Field research involves collecting original data (qualitative or quantitative) in field sites. The course will combine informal lecture and discussion with practical sessions designed to build specific skills for conducting field research in organizations. Readings include books and papers about research methodology, as well as articles that provide exemplars of field research, including both theory driven and phenomenon driven work. Specific topics covered include variance versus process models, blending qualitative and quantitative data (in one paper, one study, or one career), collecting and analyzing different kinds of data (observation interview, survey, archival), levels of analysis, construct development, and writing up field research for publication. A core aim of the course is to help students understand the contingent relationship between the nature of the research question and the field research methods used to answer it, and to use this understanding to design and carry out original field research. Course requirements include several short assignments assessing readings and a final paper designed to help students' further their own field research goals. This seminar fulfills a requirement for HBS Organizational Behavior and Management students.

*Prerequisite:* Students are required to be in or beyond their second year of study. Other students permitted by permission of the instructor.